Early Journal Content on JSTOR, Free to Anyone in the World

This article is one of nearly 500,000 scholarly works digitized and made freely available to everyone in the world by JSTOR.

Known as the Early Journal Content, this set of works include research articles, news, letters, and other writings published in more than 200 of the oldest leading academic journals. The works date from the mid-seventeenth to the early twentieth centuries.

We encourage people to read and share the Early Journal Content openly and to tell others that this resource exists. People may post this content online or redistribute in any way for non-commercial purposes.

Read more about Early Journal Content at http://about.jstor.org/participate-jstor/individuals/early-journal-content.

JSTOR is a digital library of academic journals, books, and primary source objects. JSTOR helps people discover, use, and build upon a wide range of content through a powerful research and teaching platform, and preserves this content for future generations. JSTOR is part of ITHAKA, a not-for-profit organization that also includes Ithaka S+R and Portico. For more information about JSTOR, please contact support@jstor.org.

THE RELATION BETWEEN KINDS OF STATISTICAL UNITS AND THE QUALITY OF STATISTICAL MATERIAL

SUMMARY

The word statistics may denote either the method or the material, 673 — The character of the material depends upon that of the unit, 674 — Classes of units, 675. — A. *Individual things* the quantity of which is determined by counting: (1) Natural kinds and events pertaining to natural kinds; (2) Produced kinds and produced qualities of things, 675. — B. *Mensurational units* which are applied to determine quantity without regard to individuality: (3) Physical measures (of length, capacity, etc); (4) Measures of pecuniary value, 675. — (1) J. S. Mill's conception of the natural kind, 675 — Fulness of meaning follows from the definition, 676 — Recognizability apart from definition, 677. — Counting is the typical method of determining the quantity of units of the first order, 679. — (2) Products are usually defined by way of purpose and function, 680. — Difficulties of definition, 681 — Careful attention to classification required, 681 — (3) A measure is arbitrary and is useful chiefly because standard, 682. — Individuality is ignored, 683. — Desirability of counting by size-classes, 684 — Measurement is a supplement to, not properly a substitute for, counting, 684. — (4) The importance of the pecuniary unit is due to its being the most universal of common denominators, 690. — This involves abstractness, 691. — Lack of uniformity of accounting practice a large factor in uncertainty of meaning, 692 — Economic and pecuniary statistics are not identical, 696. — The purpose of index numbers, 696. — An important way to better statistical material is to make use of the higher orders of statistical unit, so far as possible, supporting in this way all data of lower order, 697. — The development of physical and operating statistics in support of financial reports of corporations, 698. — Physical valuation in the interest of the investor, 698. — Cost-keeping a similar statistical development, 700. — Pecuniary statistics especially need supporting data of higher order, 702.

THE name statistics denotes both a method, or methods, of utilizing a certain class of facts for scientific or practical purposes, and also the facts or materials of knowledge that may be so utilized. It is with the materials accessible to exploitation by statistical methods that this paper deals. These materials are superficially distinguished as being numerical. Is it to be inferred that all numbers are statistics ? Or, to put the same question in a different way, are they all statistical material of equal grade ?

Statistics combines units into aggregates and recombines aggregates into totals. For comparison, the aggregates are then analyzed and condensed into significant averages and ratios. The quality of the product of these processes depends upon the adequacy of methods employed and upon the quality of the raw material. The quality of the material is conditioned by the character of the unit and by the completeness and correctness of the combination of the units into aggregates. In other words, the quality of statistical material varies with the adequacy of the count and with the character of the unit counted.

As regards the comparative importance of accuracy in enumeration and compilation on the one hand and of the character of the unit dealt with, on the other, it is sufficient here to affirm the equal importance of the latter and let the following pages be the evidence. So far as accuracy means exactness as distinguished from mere representativeness (if it might be assumed that we could have the second without a good deal of attention to the first) the superior importance of the character of the statistical unit, tho so little noticed, should be evident to those having acquaintance with the actual processes of statistical cumulation.

The uses to which numerical data may be put depend upon the kind of unit more fundamentally than upon anything else. The quality of a product is always conditioned or limited by its ingredients. It is, of course, fundamental to know what the unit employed means, that is, its denotation and connotation must be clear. But when all requirements of definition and conception are met, the unit may still be good or bad, and consequently, the numbers obtained more or less amenable to statistical use. It is an incident of this distinction that some divisions of statistics, in their existing state, must be given lower scientific rank than others.

The scheme of classifying statistical units here proposed is as follows: —

A. Individual things the quantity of which is determined by counting.

1. Natural kinds and events relating to natural kinds.
2. Produced kinds and produced qualities of things.

B. Mensurational units which are applied to determine quantity without regard to individuality.

3. Physical measures (of length, capacity, etc.).
4. Measures of pecuniary value.

These classes of units are arranged in the descending order of their statistical quality. In this order we shall discuss them.

The Natural Kind

The conception of a distinction of kinds as developed by J. S. Mill in his *Logic* (book I, chap. vii, § 4) is as follows: —

> There are some classes, the things contained in which differ from other things only in certain particulars which may be numbered, while others differ in more than can be numbered, more even than we need ever expect to know. Some classes have little or nothing in common to characterize them by, except precisely what is connoted by the name: white things, for example, are not distinguished

> by any common properties except whiteness; or if they are, it is only by such as are in some way dependent on, or connected with, whiteness. But a hundred generations have not exhausted the common properties of animals or of plants, of sulphur or of phosphorus. . . . If any one even chooses to say that the one classification is made by nature, the other by us for our convenience, he will be right; provided he means no more than this: Where a certain apparent difference between things (tho perhaps in itself of little moment) answers to we know not what number of other differences, pervading not only their known properties, but properties yet undiscovered, it is not optional but imperative to recognize this difference as the foundation of a specific distinction.

Demography, or population statistics, has for its principal unit the human individual, and human individuals constitute a natural kind. Other examples of natural kinds in statistics are the various raw products of the animal and vegetable world, the numbers of which are usually obtained by counting discrete units.

It is evident that where the statistical unit is a natural kind it is superior in respect of both definiteness and fulness of meaning to any that requires an artificial distinction, however well thought out. It follows that, other things equal, the highest grade of statistics is composed of numbers relating to natural kinds. Tho the best examples are the orders, species, etc., of animals and plants, any natural and therefore more or less genealogical method of distinguishing objects would give classes of the same character. "Sticks and stones," as well as animals and plants, might be divided into natural kinds, tho probably differences of behavior are so helpful in classification that unmanufactured inanimate objects would not so readily lend themselves to the differentiation of kinds as do living things. But a genealogy of atoms and molecules and even of their associations and mixtures is conceivable; and it could not fail to show natural differences of kind.

Natural classification is the opposite of mathematical. A mathematical classification might be arrived at by a development of the permutations and combinations of specified qualities which would make of each combination a species. In a natural classification, definition does not depend on a single point but oftener upon a combination of characters, some of which may be absent. To classify plants merely according to the number of pistils and stamens is in this sense "mathematical." Defining by reference to a single point and hinging classification on such definition is a thing to guard against. Distinctiveness, in the sense of recognizability, is not to be attained that way, any more than is a workable classification. Statisticians sometimes fail to distinguish between clearness of discrimination — which is best tested by promptness of recognition — and sharpness of definition.

A very great advantage of the natural kind as a unit consists in its ordinarily being recognizable without the aid of definition. It is not necessary for the enumerator to exercise great discrimination in determining whether an animal is a horse or a cow. This is not the same as saying that there will be no doubtful cases in the discrimination of natural kinds, cases where the use of carefully devised definition must be resorted to. But the discrimination of the natural kind does not ordinarily depend upon them, while in the case of an arbitrary or "mathematical" classification it is likely to.

Statistics relating to sex and race deal with differences of kind. This fact adds greatly to the significance of the results of a count, even where, as in the case of races, mixtures are of frequent occurrence and definition, therefore, not easy. The different species of domestic animals are also natural kinds. Ox-hides are, consequently, a natural kind, but shoes are not. A census of farm animals gives reliable results strictly in proportion to the care with which the enumeration is made, while a census of occupations or of manufacturing establishments may give somewhat uncertain and disputed results despite the greatest care and conscientiousness.

Variation in the size of units may be urged as an objection to the adequacy of counting as a means of determining quantity. In the case of a natural kind such variation is easily disposed of. It is ordinarily quite regularly distributed about a mean. If it be desirable to gage the character and range of such variation, the problem is one with which the statistical method is especially competent to deal. This is the nature of most statistical biology.

Variations in size may be studied as a means to the further and fuller characterization of the kind. Counting by size classes and sub-classes may be a sufficient substitute for detailed measures. Coefficients of variation for such variations as are of fairly constant character may thus be determined. When the variation, both quantitative and qualitative, about the mean is found to be very nearly constant, once these coefficients are determined, totals may convey all the information that it is necessary to obtain by actual count, the rest being easily estimated. It would seldom be necessary to specify the quantity of the sizes of grains in a bushel of wheat, or of fish in a catch, or of range cattle in a herd.

Natural kinds may undergo secular or evolutionary change. Cattle have increased greatly in weight in the past few centuries. This is an important statistical element which might seriously affect a comparison. But the presence of such a change is easily ascertained.

To make proper allowance for its influence is not difficult. Secular change in the character of the statistical unit is one of the least troublesome phases of definiteness.

Counting is the usual method of obtaining statistics of natural kinds. Measuring, tho sometimes a practical convenience, is never a logical necessity. Counting will tell more than mere measuring.

Differences in the degree of some quality of a natural kind partake of the statistical character of the unit to which they pertain. Age is an example. That it is measured in terms of astronomical periods is logically an accident. Ages represent differences in the degree or stage of development of a group of natural qualities, called, according to the varying stages, youth, maturity, senility, etc.

Births and deaths are events relating to natural kinds and they have a corresponding statistical standing. Thus vital statistics in general have a unit of the first order. But the status of a marriage is somewhat different. At least as legally defined, it is as much artificial as natural.

To summarize the conclusions drawn from the foregoing: Statistics of natural kinds are superior to such as are based upon some other sort of unit, both negatively and positively. Misunderstandings of the informant, of the enumerator, and of the compiler should be at a minimum in the case of such a unit. Instruction in niceties of definition is seldom necessary. And this ready recognizability is no accident; it results from the nature of things. Positively, also, such statistics have fuller meaning and may at any time develop an unexpected significance. A division into natural kinds is more concrete than one requiring abstract definition. Relations with other objects are clearest where the terms are ordinarily concrete or discrete kinds. Much may be made of statistics of natural kinds even where there has been no careful attention to classification.

Products as Statistical Units

In calling the second order of statistical units "produced" objects we bring in the idea of economic production. Economic goods are the result of modifications of natural materials for human uses and purposes. The materials may still have the properties of natural kinds. But the statistics of such things will relate primarily to the distinctly produced qualities. Classes of these are not fixed and objectively definite. The purpose and function of the same article may vary, and physically different goods may, on the other hand, be made to serve the same purpose. A door is not a door when it has been converted into a table top. But it is not so easy to say when a street car ceases by reason of decrepitude to be a car. Corn is not fodder when it becomes fuel. If a chair is something to sit in, what is a stepladder chair? Personal idiosyncrasy as well as human reason may be a factor in such classifications.

The foregoing illustrations suggest the fundamental difference between units of the first order and those of the second order. Among the latter, definition and classification hinge mainly upon function. Natural kinds are defined otherwise. It is significant that the biologist finds functional characters of little or no use for the purpose of classification.

Manufactured commodities and instruments and the produced qualities of natural and other objects comprehend virtually the whole of the second order of units. The drawing of the line between such objects and natural kinds is not always so easy as might appear, but that need not detain us.

With this sort of unit difficulties of definition do not, as with natural kinds, amount to less in practice than in theory. An important group of such difficulties consists of those resulting from a compound purpose — the case of a tool that will do several things equally well, or a wage-earner who has two occupations. Difficulties of definition are familiar in relation to numbers employed by a particular concern or in a particular occupation. Complications due to time lost, part-time employees, and subsidiary occupations, are ghosts that refuse to be laid. Apparatus in use or available will be returned variously until the use and the degree of need of reserves for repairs and emergencies becomes more nearly determinate. It is not an accident that these examples, tho there was no such intention, come from industrial statistics. Here much more depends upon good classification, or rather upon careful attention to classification and definition, than is the case in dealing with natural kinds.

Freight and passenger cars of the railways are characteristic made kinds, with a good deal of recognizability, but sometimes also occasioning much perplexity in border cases. How shall cabooses, how combination express and mail cars be classified ? One street railway has been known to return the little-used private car of its president as a freight car. Difficulties are sure to arise where classification depends ultimately upon purpose, even tho incidental details of physical construction ordinarily help — perhaps in the end only to increase the doubts and difficulties when they do arise. It is the purpose of the maker and of the user not the observer's notion of what is suitable, that is referred to, hence the criterion is comparatively objective. To define by the purpose of the maker is one way to make the required definitions clear and easily applied. This should be done, even at the cost of some degree of arbitrariness.

Produced qualities as well as produced objects have the characteristics of this second sort of unit. Social status in general, for example conjugal condition, probably belongs here. Occupations, and with them statistics of wages, clearly belong here. Wages really describe a produced quality of the person in the occupation, or the degree of such a quality. Hence wage statistics are not to be classed under the value unit, where they appear to belong. They are attached to a definite object, a human being, and afford knowledge of a certain produced quality.

It is probable that, as time goes on, this kind of statistical unit will tend to improve in character, quite apart from any statistical interest in such improvement. The modern tendency towards standardization of instruments and products is of increasing importance as markets become larger and productive processes more complex. Staple and standardized commodities are better suited to such conditions. Commercial variations in a given commodity that are not dishonest will tend to be fewer because the consumer will feel himself less able to cope with complexities. Statisticians and administrative officers will doubtless discover and apply means of preventing dishonest variations in commercial units. It does not so much matter what the unit is, provided it be fairly constant. That matters more and more, in practical ways as well as scientifically.

There are units which are intermediate in character between the first and second orders. Such is apparently the family. In itself natural — at least there is a natural family — so many of its functions and traits are artificial or economic that the United States Census, perforce, defines it artificially. Yet inferences as regards the natural family derived from such statistics are not entirely bad.

The city also, so far as it is a statistical unit, has a similar intermediate character. Inequality of size would not be an objection to its use, tho any exploitation of such figures could scarcely fail to take account of this. If a fully natural kind, its definition would not, of course, hinge upon corporate geographical boundaries. But cities are rather too few in number for the "trees" to lose their individuality in the "forest," hence there is less need of, and less opportunity for, aggregate or statistical treatment of the city as such.

The business corporation is legally and in some other respects analogous to the city. Here there is certainly a "forest." There is no reason why we should not have true statistics of corporations. The unit is not perfect. It is in the main a produced unit, and the size and internal organization of corporations vary greatly. It is not entirely a produced unit, however, since it is a group of persons and has most of the qualities of the individuals who compose and direct it.

Whether nationality is a produced quality or a phase of a natural difference in kind may be doubted. It is a border case. A Frenchman may be of Teutonic stock, but he is French if his language, culture, and traditional sympathies are French. The characteristic element in statistics of immigration is usually nationality, hence such statistics are more of the second order than of any other.

Statistics of manufactured objects are gathered mainly by means of enumeration and their class or descriptive name will ordinarily tell much about them. But it will not always tell enough, even where much discrimination is exercised in determining what is to be counted and even when, also, the classes of objects are subdivided in the count. Hence a specially significant unit, one that is indirectly a unit of capacity, may sometimes be employed as a basis for addition and comparison. Thus statistics of cotton-spinning manufacture take as their unit the spindle. In some cases odd units may be reduced to a common denominator, as in the case of the standard 500 pound bale of cotton-ginning statistics.

Sometimes counting produced objects quite fails of its purpose and resort is had to measurement. Hence statistics of capacity in combination with number, examples being the tonnage of ships, the capacity of grain elevators and of engines for power production. This case is transitional to the next class of statistics, where the unit is primarily a measure of size or capacity instead of a concrete object. Counting by size-classes or grades, however, may sometimes be preferable to measurement.

Physical Measurement Units

Physical measurement units compose the third kind of statistical unit. Examples of physical measurement units are the ordinary measures of length, of cubical capacity, and of weight, and measures of energy like the horse power and the kilowatt hour.

The size of such a unit is the result of accident and convention. The length of the foot of some king is as good as a decimal fraction of the miscalculated circumference of the earth. The horse power may as well have a merely arbitrary as any other relation to the power of an average horse. There is no particular reason why the yard or the pound should count for one, except convention and convenience. There should be some common and familiar standard. What it is, signifies little. Similarly it is necessary to have a rule of the road, but whether it says that one shall turn to the right or to the left does not matter. That the unit be accepted as standard is all that is necessary. There is a marked contrast between this situation and the way in which the statistical unit is determined for a natural kind. There we find nothing arbitrary or conventional.

But, tho the fact that units are merely conventional does not matter, there should be no variability or ambiguity in the convention. Our abominable English weights and measures include several different kinds of pounds, tons, quarts. Quantities reported in tons are only presumptively known unless the unit is described every time it is given. The most definite schedules are likely to be filled carelessly.

Engineers have been too ready to accept a situation as regards their peculiar technical terms which is not much better. Statistics of rated capacity of power equipment suffer considerably from lack of standardization. But the engineering societies are now attending to such matters. Mention may be made especially of the American Institute of Electrical Engineers. The work of the National Bureau of Standards insures steady improvement in metrology generally.

The very fact that the unit is a measure more or less arbitrarily arrived at, and incidentally become customary or standard, suggests its limitations. Measuring ignores individuality and disregards all but one of the qualities of the objects measured. In statistics of the third order we no longer count the members of a kind and in naming also describe them. The significant relations of a measurement unit to things in general must, of necessity, be narrow because unilateral. They take account of only one thing at a time. A particular steam engine has a determinate relation to the men who run it and to the machines for which it furnishes power, but "ten horse power" has no such relation to other things. It tells something about an engine, but does not mean the same for a ten horse power engine as for one of a hundred horse power. It is more important to know the number of engines and their size (by classes) than their aggregate power. But we sometimes have to be content with horse power alone.

Measurements always involve abstraction. A measurement gives length, or cubical contents, or specific gravity, any one of which abstracts from all but a single quality or relation. Diverse measurements may give several such facts about a series of objects, but the results are still abstract. If we are to learn about the real concrete things it must be by way of supplementary or collateral information, not always statistical in character. Capacity may be measured, disregarding shape; weight, disregarding material; the figures thus limit themselves. The counting of discrete natural objects, on the other hand, gives numerical knowledge which can be supplemented to any desired extent by reference to natural qualities.

When objects are both measured and counted, which is to be regarded as the primary datum? On the whole the counted object should be primary because of its superior character as a unit, tho the greater importance of the measurement unit may sometimes outweigh this consideration. It is well, whenever possible, to have both sorts of quantity and also to have the two clearly related, *i. e.*, the numbers by sizes or by size-classes.

Sometimes what appears to be measurement is in effect only a quicker way of counting. It may be more convenient to determine the number of new coins in a package by weight than by tale. But the result is not less a number of coins. Commodities are often weighed merely because this is the best way to arrive at their amount. Bushels are used to measure grain. The object so measured is a natural kind composed of individual things. Relations to other things, for example to nutritive value and tastiness, are not at all left out of account. In these cases measures are not resorted to because of diversity of type and of size among the objects, as in the case of produced articles such as generating engines. Natural kinds seldom require this, even where we are directly interested in their size, for they vary regularly about a representative mean. But the size of some marketable products is so largely produced that it has almost ceased to be a natural property of the kind; hence, there is a tendency to determine quantities of such things by weight, even when the article is sold in its natural state. If it is becoming true, however, that eggs should be sold by weight instead of by count, this is chiefly because of the work of the breeder in developing marked varieties of the domestic fowl. The size of the egg is coming to depend on the breeder's art.

Measurement aggregates may consist of one continuous quantity or a homogeneous mass, or they may relate to a miscellaneous lot of individual things which are varying multiples of the measurement unit. The latter case seems to be the more frequent in statistics, perhaps because the more evolved and individualized things, whether made such by nature or by man, that is whether biological species or manufactured articles, are of more interest to us. But it is probably with the former class of materials that measurement begins, later extending to objects that might also be counted. The homogeneous material may, of course, be a natural kind, but not of the most interesting order. So far as it is such, a measured amount of it, of water or sand, for example, may have some of the advantages of the first order of unit — some only, however, because such unintegrated matter too readily mixes and mingles with other things.

Accuracy of measurement deserves notice in this connection by way of distinction from accuracy of counting, and also for comparison with it. The definition of the measurement unit offers practically no difficulties. The number of units to be recorded usually involves some mathematical computation upon the basis of comparison with a standard measure, or it may be obtained by estimation upon a more or less objective basis. There is not so much difference between measuring and objectively estimating quantities as one might easily assume. The view of an expert with the time and facilities for applying objective tests to check his results is worth much more here than it is in the field of counted units. The method of sampling can, in this case, if intelligently used, be adequate for most statistical purposes. The way in which several measurements are employed to check each other by the physicist and mathematician and are assumed to be inevitably more or less approximate, illustrates the necessary inexactness of human measurement and also suggests that “ mathematical exactness ” may often be practically unimportant as well as unattainable. If we dealt with the objectively checked estimates of impartial experts as the mathematical physicist deals with quantitative observations, we might find them very nearly as good for most of the purposes of statistics as " exact " measurements.

It sometimes happens that the reduction of odd articles to a common measurement standard is desirable. Provided the material contained is homogeneous, little is lost. But caution in the use of this short-cut is imperative. The day's work, for example, is not merely so many hours of work. A 16-hour day is by no means equal to two 8-hour days; nor is a 48-hours-a-week schedule the same regardless of how or when the time is put in. Sometimes administrative regard for facility of enumeration may cause a too easy acceptance of the measurement unit in such cases, when varieties and their relative importance are quite as interesting as totals and averages.

Some statistical units in common use are compound. This is especially true of units of physical capacity and performance, that is, of such as have the characteristics of our third class. Such a unit usually takes up whatever disadvantages pertain to both its terms. The foot-pound is only apparently such; its compound name resulting from the attempt to make concrete an abstract conception. Foot-pounds per hour is truly compound; hence horse power is compound. The car-mile and car-hour are compound. The vagueness that attaches to these units follows from the character of the basic unit, which is of the second class. It is important to know the size of the car and whether trailers are used. But the grade and curvature of the road also affect the significance of such a unit. If the idea is to measure the potential service performed by a passenger car, the seat mile is more to the purpose. One term of this unit is of the first order. The ton-mile unit is of the third order as regards both its terms. But data for ton mileage and seat mileage do not make it possible to dispense with car mileage.

Comparability is the fundamental desideratum of statistical data. There is no difficulty about it with natural kinds. There is a good deal with produced objects. The difficulty is apparently met by the use of measurement units, but the solution is often apparent only. The additive quality is secured, but perhaps at the cost of losing trace of important relations to other things. Hence measurement as a means of obtaining a total aggregate amount is to be considered a supplement to, not a substitute for, counting, whenever the quantities dealt with are composed of discrete natural objects.

The Pecuniary Unit

The fourth order of statistical unit is the unit of commercial value. It is possible that some other than the monetary measure might be found that could serve as a basis for value statistics. In fact, however, the only sort of unit we use is something reducible to the dollar. Whether any and all statistics of this fourth class should be called financial, or financial and commercial, as seems to be the tendency, is questionable. But financial statistics are representative of the class.

One reason why the pecuniary unit bulks so large in common statistics, including business statistics, is because it is the most universal of common denominators. There is not much of practical interest in the world that does not, at least occasionally, have a pecuniary value assigned to it. Nor is the tendency to attribute to all things a pecuniary value so reprehensible as is often alleged. The error, which is a moral as well as intellectual error, consists not in the extended application of the pecuniary measure, but in supposing that it has an intensiveness, comprehensiveness and independence of meaning which is foreign to its nature. Just because of the general applicability of the dollar as a common denominator, its definable content and meaning must be small. All measurement units suffer by reason of their abstractness. The most universal of common denominators, despite its fundamental interest for every human being living in an exchange economy, will naturally suffer most in this way.

In citing examples of the fourth order of statistics we can hardly refrain from commenting on their quality. Statistical inferiority is characteristic of numerical data based upon the pecuniary unit. A review of examples of such statistics becomes a process of learning why this is so. Hence this judgment may as well be put at the front of this section. It applies with full force, however, only when the pecuniary unit is used by itself and as standing on its own bottom.

Financial reports are the sources of most pecuniary statistics. Our statistics of railways, in the main pecuniary, are among the best examples of the class. An authoritatively prescribed uniform system of accounts will make them vastly better than they were before. Statistics of assessed valuation and of taxation are a less developed variety of the fourth order. But, if we wish to prove that we are economically better off than were our forefathers a hundred years ago, it would be sounder to argue less from the increase in *per capita* wealth than from the application of invention to promote abundance and variety of goods. Statistics of export and import trade are almost altogether pecuniary as to their unit, largely because of a possibly too great regard for administrative convenience. Hence it is not easy to determine just what is the status and recent tendency of British trade or just what our own trade balance signifies. Aside from disturbing factors peculiar to international trade, the level of prices changes with time and place and with it the meaning of the pecuniary unit. Finally, our statistics of capital and capitalization are most completely pecuniary. Despite great practical interest in them, all such data are, as compared with demographic statistics, scientifically barren. The almost exclusively practical interest of their compilation is no sufficient explanation. On the contrary, this should provide richer and better material for analysis by the scientifically disposed.

One reason why pecuniary statistics are so generally unreliable as to their comparability, and therefore as to their significance, is because they are usually derived from books of account. Accounting theories and practices are anything but uniform as regards their treatment of particular items. Indeed, less elasticity of method and usage would often serve less well the purposes of business men. Accountants, for their lack of breadth of view, deserve part of the blame. But the public interest in the intelligibility of financial statements is gradually establishing authorities to prescribe and compel uniformity in such matters. Until uniform systems of accounts have been in effect long enough to insure their smooth and accurate working, our financial compilations will fall much short of being entitled to the standing that should be connoted by the name statistics. Accounting entries ought to have constant significance and comparability. The mere accountant has been too much inclined to make his own financial balances the *ne plus ultra* of his work.

The investor is likely to form his opinion of a stock on the basis of what it pays in dividends. But if he is alert he will consider rather the net income earned from year to year. Tho much speculation has just so superficial a basis, a dealer in stocks will go deeper. The investment expert of a broker or banker will go as far into a company's reports as there are data. He will, if possible, analyze the physical statistics. That he does not give more attention to such matters may be explained by the usual absence of the necessary information. Even the expert, however, is likely to give the investor chiefly the tabulations of gross and net income to which he is accustomed rather than educate him up to some understanding of the physical basis of profits.

If profits are really there, the basis is physical as well as economic. But the basis of net income as it stands in a company's report may be neither physical nor economic. With accounting methods not yet fixed and with the management unchecked by statistical indices of physical condition, it is quite possible, especially through maintenance accounts, to juggle with net income in ways that no superficial analysis will disclose. Depreciation can be charged only on an accrued or estimated basis and we have as yet no sufficient experience to check the basis, tho we have had plenty of experience of the arbitrariness and manipulation of maintenance charges. All this relates to property operated. Intercorporate relations and resulting elusive "other income" offer another handle for manipulation. The line between capital expenditures and operating expenses is in its nature indefinite. That between expenses of the period and of other periods (or years) is equally so, especially on account of depreciation. Hence the necessity of honesty and — as a guaranty of honesty and competence — full data to enable the outsider to form an opinion. "Accrued" income may often be more properly described as contingent. Hence the "actually paid" basis of the federal corporation tax. The government can wait for the long run. But the outside investor may be "cleaned out" long before the consummation of long-term changes whose probable effect he is not permitted to see.

Merely financial reports are not adequate to the use to be made of them. It is of the essence of accounts that they cannot be so. "Net income" is as much a result of how the books are kept as of profitable business transactions. It should not by itself be the test of solvency but should be supplemented by full operating and physical statistics.

The inferiority of the fourth order of statistics, however, is due to more fundamental causes than arbitrariness of accounting practice. The dollar is intrinsically inferior as a statistical unit. The physical measurement unit as well as the pecuniary unit is the expression of a relation and not discrete; but the former can easily be so conceived while the latter cannot. The dollar unit is so completely abstract as to be not even imaginable. It is a measure of "power in exchange." But the dollar's power in exchange does not mean the same thing to any two persons. Even if subjected to the test of actual exchange and determined by market conditions the purchasing power of a dollar fluctuates, hence estimation must be resorted to and allowance made for time and place. Even then much depends on "whose ox is gored." Financial data usually have the character of interested estimates. Differences in valuation are not mere variations around a mean and are not moderate. Hence they are not easily dealt with or lightly to be disregarded. Reducing to a fine gold equivalent is merely a way of getting around diversity of currency. Changes in rates of exchange are a minor adjustment, for custom and "psychology" are factors in price. The symbol is familiar and constant. That for which it stands changes in ways not easily determinable, certainly not to be simply ignored.

The method of index numbers — itself a nice problem in statistics — largely meets such difficulties with the value unit as are due to changes in time. Logically, however, the index should change with every purpose for which statistics of the fourth order are used. The proportionate weights of the prices of particular articles which are determined by the consumption of wage-earners in a manufacturing community seem to be of greatest use. But they are not therefore generally adapted to all purposes, even tho adjustment of the weights cannot be alleged to be indispensable.

The financial statistics showered upon us do not have the definiteness of meaning that statistics may be expected to and *ought* to have. Efforts should be made to close the existing great gap in scientific character between financial and demographic statistics. For taxation and financial administration and for the public control of corporate management, adequate statistical data will become more and more necessary. This refers to the need of governmental action. Purely private interests (if there be such) are also involved. Many large corporations are coming to feel that they ought to make the knowledge of their operations accessible to the public, at least to their actual and possible stockholders. The statistician thus tends to be a necessary aid in the conduct of every large corporate enterprise.

A suggestion of the direction which improvement will take is contained in the rather obvious proposition that economic and pecuniary statistics are not necessarily the same. The backward state of economic statistics, as compared with demographic, is doubtless largely due to a too ready acceptance of figures of value as a satisfactory answer to economic questions. But economic statistics need not be so superficial. They will, when fairly complete, doubtless contain values, but will not be composed of them. Statistics of manufactures and trade ought to present quantities and kinds of products as well as values or prices. Our agricultural statisticians do this very largely, tho they too sometimes seem to prefer to detach the values from all other numbers relating to their data. Even where the interest is purely commercial or financial the trend is towards more adequate figures. We are told of tons of rails produced as well as of the value of the output. It would be still better to have the kinds specified, *i. e.*, the weight per yard and the material, perhaps also the shape, — things which, it should be observed, are seldom necessary in statistics of natural kinds. Our numerous practical " statistical " manuals designed for investors need to pay more attention to such possibilities. But the large corporation must first appreciate the importance of compiling and making public more adequate data.

Index numbers have been referred to above as affording a way out of some of the difficulties of statistics resting upon the pecuniary unit. This is a way to get beyond pecuniary value and down to concrete goods. We may thus learn what the dollar means in terms of necessary articles of consumption. The device is practically the determination of an equation between money and commodities. The principle of this solution ought to be more generally applied.

A Way to Better Statistical Material, especially to Better Economic Statistics

While it is of some scientific interest to distinguish different orders of statistical units and of statistics, what the distinction is really worth depends upon what purposes it may serve. For statistical practice the important question is: How can the inferior orders of statistics be improved ?

Statistics first achieved scientific standing in the field of vital statistics and demography, that is, in a general way, among natural kinds. Whether as an extension of the view-point of "political arithmetic" or as an intensive development, statistics that are something more than mere numbers have come rather later in other fields. Among the divisions of our federal census the volumes on population are, as statistics, the best; those on agriculture (the units here also being chiefly natural kinds) rank next; and last or lowest in quality are the statistics of manufactures, the reason being that the unit is too largely pecuniary.[1] One might well infer that our statistics of manufactures should be made less dependent on the pecuniary unit. This is not meant to imply that our census officials are unaware of the need.

---

[1] The defects of the volumes on vital statistics — the statistics obtained by enumeration as distinguished from the recent compilations of mortality registration records — do not vitiate the comparison because their condition is due mainly to an accident in the development of our statistical administration.

The way to improve pecuniary statistics is to relate values and prices, present and past, to the things to which they pertain, or rather to the statistics of those things. Pecuniary statistics should be, so far as possible, either accessory to or supported by statistics of higher orders. An amount paid in wages immediately provokes a question as to the number and kind of wage earners, how long they worked, and whether on full time or part time. It is not enough to know the value of our exports of a particular class of commodities, tho the difficulty of defining varieties does in part excuse the shortcoming. Railroads want to know the relation of expenditures to work done, hence ratios of cost and of traffic to various operating units, expenses and revenues per train-mile, ton-mile, etc. New uses of such units and ratios are constantly appearing. Detailed power and wage statistics serve the same purpose and will soon be considered a necessary part of the classified and analyzed records of any large private corporation. It is a need of our mental nature that values be predicated of definite things. As the workings of our minds are pragmatic, this need points to a function.

Especially statistics of capitalization need to be related to the statistics of capital or of the means of production upon the basis of which securities are issued. This happens to be a degree removed from the direct interest of active business men, hence "neglect" does not sufficiently describe their attitude in the matter. Some of the uses to which physical valuation might be put may reasonably be objected to, but to the thing itself there is no valid objection. The investor should be put in position to know his property. The only way in which such knowledge can be conveyed to his understanding in the case of a large corporation is by statistical description. This is the foundation and essence of physical valuation. A corporation cannot properly have secrets. Especially in relation to its stockholders, its duty is not merely the negative one of abstaining from withholding information. Even if the corporate net income be the basis upon which the investor buys the stock, its stability as well as its amount is important and the probable stability of the income depends chiefly upon an adequate physical basis.

Economists of a certain tendency share with business men the blame for inadequate notions of how financial statistics ought to be looked upon, at least as regards the phase of them that relates to capitalization. The capital value concept is useful and important. The attempt to wash out the distinction between capital as a sum of value and capital as concrete means of production, or what amounts to this, to ignore the latter meaning, is very much the opposite of useful. The dominance of the corporate form of organization makes it important not to destroy, but to develop the idea of capital as concrete means of production. The small individual entrepreneur does not know less of what he possesses by reason of the fact that he reduces all to terms of value as a common denominator. The stockholder of a large corporation is differently situated. He can know what he possesses and how it is being used only by means of adequate physical and operating statistics, including data of performance, duly reported to him. Mere balance-sheet and income statements are not sufficient. Here is where the capital-value theorists may have occasioned some distraction of attention from real needs. With adequate inventories of extant properties, capital assets could not be made to include dead horses, tracks removed, obsolete and unused implements, all and sundry covered by combination and by temporary monopoly or other advantage, without security for the continuance of such advantage and thus with loss to the investor. Uniform accounting systems should largely take care of this need as well as of the proper treatment of accounts in the customary narrower sense, as distinguished from records.

Accounts should, in general, be supplemented by records or statistics, tho the word so used sounds pretentious. The inventory made once a year, or oftener in the case of a mercantile stock, should not be dispensed with in the case of fixed capital. Tho it need not be made so often, we know that fixed capital also may disappear in ways not accounted for. Corporations have their "storekeepers" who carefully check and record whatever materials and supplies they receive and issue. Should not there be records also of the more valuable tangible assets? As statistics constitute the state's means of knowing itself, so the only available means by which a large corporation can know itself are likewise statistics, both physical or static statistics and operating statistics. So far as size is the compelling factor, there are modern business corporations quite as large as were important states when statistics first began to be.

So-called cost accounting, or cost keeping, occupies much the same position in relation to operating statistics as does physical valuation in relation to the knowledge and quantitative estimation of fixed investment. It is an attempt to show the connection between fundamental physical facts — quantity and kinds of raw materials, labor time of various sorts, space and power required — and the quantity and value of products. It is essentially statistics rather than accounting, since in it numerical comparisons are more important than the registering and balancing of pecuniary obligations. The development of cost-keeping in the business world is a phase of the increasing importance of statistics.

Any considerable acquaintance with compilations of numbers in tabular form should make it clear that not all figures are statistics, nor all figurers statisticians. Economists without special statistical interest or training have probably contributed to the general misconception of statistics as merely masses of figures. The large corporation is now coming to feel it desirable to have a statistician, tho he is still likely to be regarded as a special kind of clerk, one who is "quick at figures." Sometime it will be recognized that neither an accountant's nor an engineer's training specially qualifies for statistical work. The statistician should be grounded in demography and familiar with the statistical methods used in dealing with natural kinds, no matter how "practical" his later work is to be. It is true also that he cannot have too realistic a knowledge of his special field. The differentiation between the accountant and the statistician will probably come to mean that the latter is specially competent to deal with physical things in their numerical and quantitative aspects by means of methods suggested by a study of the higher orders of statistics.

Figures are symbols. They serve only to connect our thoughts with things. Not all statistics do this with equal directness and sureness. It holds especially of figures based upon the commercial unit of value that the meaning is not self-evident. The measure of value is less intelligible than physical measures. Statistics based upon measurement units are, in general, less good than counts, whenever the kinds counted are adequately classed and of well-defined character.

Realistic comprehension is scarcely possible with value measures unless they are definitely related to things. Things valued should be also defined and counted, or at least physically measured. The statistician must see the quantitative relations of the numbered things, not the mere symbols or numbers. As statistics becomes of greater recognized importance and as its methods are applied to more and more branches of knowledge — being thus of doubly increased significance — the greater will be the need to attend to such considerations as are mentioned in this paper. We shall supplement statistics of produced kinds with measurements, instead of beginning and so largely ending with the most abstract of measurement units, thus getting the cart before the horse. Without implying that the other orders cannot sometimes be helped by extending quantitative enumerations and analyses to cover value, it is clear that the latter kind of statistics has much less to give than to receive.

To say that pecuniary statistics are less good than pecuniary statistics plus other kinds of related figures supporting them, is an affirmation of the obvious. Yet this fact needs emphasis. Pecuniary statistics greatly need such support. Nor is it quite true that the rule works equally well the other way. The pecuniary unit is the weak member which ought as little as possible to be left to stand alone. Stated values are seldom disinterested and their basis is never quite objective. We need always to know kinds and grades and their numbers or quantities, along with values. The foundation of all statistics should be in natural kinds or in the best obtainable substitutes for them.

G. P. Watkins.

Public Service Commission,
New York City.